Image-based Onion Disease (Purple Blotch) Detection using Deep Convolutional Neural Network

Muhammad Ahmed Zaki[1], Sanam Narejo[2*]

Muhammad Ahsan[3], Sammer Zai[4], Muhammad Rizwan Anjum[5], Naseer u Din[6]

Department of Computer System Engineering, Mehran University of Engineering and Technology, Jamshoro, Pakistan[1, 2, 3, 4, 6]

Department of Electronic Engineering, The Islamia University of Bahawalpur, Pakistan[5]

Abstract—**Agriculture on earth is the biggest need for human sustenance. Over years, many farming methods and components have become computerized to guarantee quicker production with higher quality. Because of the enlarged demand in the farming industry, agricultural produce must be cultivated using an efficient process. Onion (Allium cepa L.) is an economically valuable crop and is the second-largest vegetable crop in the world. The spread of various diseases highly affected the production of the onion crop. One of the serious and most common diseases of onion worldwide is purple blotch. To compensate for a limited amount of training dataset of healthy and infected onion crops, the proposed method employs a pre-trained enhanced InceptionV3 model. The proposed model detects onion disease (purple blotch) from images by recognizing the abnormalities caused by the disease. The suggested approach achieves a classification accuracy of 85.47% in recognizing the disease. This research investigates a novel approach for the rapid and accurate diagnosis of plant/crop diseases, laying the theoretical foundation for the use of deep learning in agricultural information.**

Keywords—Disease detection; disease classification; artificial intelligence; inceptionv3; deep convolutional neural network

I. INTRODUCTION

Farming production is the utmost important sector in many countries for contributing to domestic incomes [1]. It plays a significant role in stimulating the economy of agricultural countries like India, Bangladesh, Indonesia, Iran, Turkey, Afghanistan, and Pakistan which depend heavily on crop quality and growth [2]. Productivity in agriculture is the backbone of a country's economy. With a population of more than 220 million people, Pakistan is the sixth most populated country on earth. Approximately 64% of the population, around 140.8 million lives in rural areas [3]. Agriculture plays a crucial role to guarantee food safety for a relatively enormous and increasing population. In the agriculture market, vegetables are of prime importance. The share of vegetable production in Punjab is 95.4% for potatoes, 19.9% for onions, 21.1% for chilies, and 15.9% for tomatoes [4]. Punjab is by far the largest province for the farming industry in Pakistan. The total area under agriculture in Punjab is 40.55 million acres, which makes 71.53% of Pakistan's total agricultural area [5, 6]. Sindh, in particular for tomatoes, chilies [7], and bananas [8], is the second most important province for agricultural production. Onion prices have been falling in the global market in 2020 since February, as logistical problems from covid19 [9] hampered trade and demand in major producing countries, India and China remained lukewarm. Though demand in America has improved, the recent onion outbreak may have an impact on the United State (US) and Mexico export situation. Mexico's onion harvest is divided into two seasons, the spring season lasts from March to June and the fall season from August to January. The volume of onion exports to the US, its main importing destination has increased significantly over past years with a nearly 80% increase from April to July compared to the same period last year. Since the outbreak began demand for onion in the US has been stable with the bulk of it coming from retailers. Onion cultivation in the US appears to be quite consistent with no weather anomalies. According to the onion business the coronavirus had little influence on US onion shipments to Asian countries because the season extends from August to January the remaining seasons are supplied by New Zealand. With the increasing cargo charges moving the product across the country could be difficult but the foodservice business is warily hopeful about reopening which might handle the onion that began packing in August. During the early stages of covid19, trade flows were restricted due to logistical issues that made it difficult to export Chinese goods and lockdowns in Asian countries that slowed down exports. However, the trade situation has now returned to normal.

In Pakistan, a brief overview of the main crops is highlighted in Table I. Onion (Allium cepa L.) is the world's second-largest vegetable crop with an international supply of approximately 74.25 million tons [10] and is commercially the most important crop of the Alliaceae family [11]. Onion is the main relish commonly used throughout the year in all homes [12]. Recent research has shown that dietary onions can help prevent heatstroke and other disorders [13]. The onion bulb [14] is abundant in carbohydrates, calcium, and phosphorus.

Table II indicates the composition of the onion bulb. Food historians agree that onion originated in central Asia 5000 years ago, some suggest onion was initially grown in Iran and West Pakistan [15, 16]. Since onion has grown wild in different areas it has probably been eaten for multiple years all over the world [17]. Onion may be one of the initially cultivated crops since they were simple to harvest, less fragile, transportable, grown in a diversity of climates and soils than other crops [18] as shown in Fig. 1.

*Corresponding Author

TABLE I.   CASH CROPS OF PAKISTAN

Major Crops	Area (HA)	Production (Tons)	Global Ranking (Value)
Cotton seed	2805700	4071400	3rd
Sugarcane	1128800	63749900	5th
Mangoes, Guavas	171289	1658562	6th
Wheat	8686602	24211400	8th
Onions	125900	1660800	8th
Rice	2789200	6798100	13th
Maize	1168490	4944210	22nd

TABLE II.   COMPOSITION OF ONION BULB

Carbohydrates	11g	Calcium	0.2g
Moisture	87 g	Vitamin C	11 mg
Protein	1.3 g	Phosphorus	0.05 µg
Thiamine	80 µg	Riboflavin	0.01 mg
Fiber	0.65 g	Iron	0.8 mg
Minerals	0.45 g	Nicotinic acid	0.4 mg

Fig. 1.   Fresh Onion.

Besides, onion was beneficial for human life to be sustained. Onion avoided thirst when food gets scarce could be dried and stored for later consumption. Onion helps lower blood pressure [19], lessen the occurrence of cancer [20], insect bite antiseptic, anti-diabetic, anti-aging, and stimulate hair regrowth [21]. The total world production of onions is 742.5 million tons with Pakistan ranked 7th [22] as presented in Table III.

TABLE III.   WORLD TOTAL ONION PRODUCTION

S No.	Countries	Million tons
1	China	205.08
2	India	133.72
3	USA	33.22
4	Egypt	22.09
5	Iran	19.24
6	Turkey	19.01
7	Pakistan	17.02
8	Brazil	15.57
9	Russia	15.37
10	Republic of Korea	14.12

TABLE IV.   STATUS OF ONION (PROVINCE-WISE) 2006-07

Province	Area (HA)	% share	Production (Tones)	% share	Average Yield t/ha
Punjab	35.6	27	315.7	17	8.9
KPK	12.0	9	206.1	11	17.2
Sindh	45.6	35	593.1	33	13.0
Balochistan	38.2	29	701.6	39	18.4
Pakistan	131.4		1816.5		13.8

Onions are generally grown in the field with a 10-12 cm seedling gap and 25-30 cm out in flatbed rows. To generate huge bulbs, space the crops 10-15 cm out in a row. Maintain onion weeds free by hoeing and shallow cultivation. Planting seeds directly in the field where the crop will mature, planting in a seedbed from which the plants will be moved subsequently to the field, and planting sets are the three most distinct techniques of planting onion. These sets can be bought or cultivated from seed by the grower. For early production, the transplanting approach is widely used. Onion is grown commercially on an area of 131.4 thousand hectares, producing 1.8 million tons [23]. The onion status (province-wise) for 2006-07 is shown in Table IV.

22 districts account for over 77% of Pakistan's total onion production [24]. In the world 66 syndromes including 38 fungal, 10 bacterial, 3 biological, 6 nematodes, 1 sponging plant, and 1 phytoplasmal ailment, 7 various infections and ailments are affecting onion. Numerous serious infections such as soil-borne ailments are becoming extensive and sufficiently extreme to restrict onion production worldwide. Fungi and bacteria can cause a variety of diseases in onion. Crop rotations, climate, storage, drying conditions, and disease control measures disease intensity. Botrytis blight, botrytis neck rot, black mold rot, pink root, botrytis bulb rot, bacterial soft rot, iris yellow spot virus, fusarium basal rot, purple blotch, translucent scale, and downy mildew are onion diseases worldwide [25]. Many onion diseases start on crops in the field and progress to the bulbs during transportation and storage. Understanding that postharvest diseases arise in the field is the first step toward effective control. Controlling onion diseases require suitable cultural practices such as removal of contaminated onion, crop rotations, culls, debris, and cultivar range [26]. Some common onion crop diseases found worldwide are presented in Table V and Fig. 2.

Following are the main contributions of this paper:

1) We demonstrated the feasibility of our approach by creating a dataset of 1000 images of onion crops.

2) Proposed an enhanced InceptionV3 classification model, which is used to identify purple blotch affected onion based on their images.

3) To improve classification accuracy, various pre-processing and training techniques are used.

4) The proposed model is created to aid in more accurate and efficient detection of onion disease (purple blotch).

TABLE V.   COMMON DISEASES OF ONION CROP FOUND WORLDWIDE

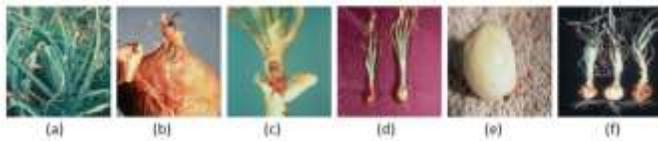| Diseases	Caused By	Pathogen
Botrytis Blight	Fungus	Botrytis Squamosa
Botrytis Neck Rot	Fungus	Botrytis Allii
Purple Blotch	Spores Of The Fungus	Alternaria Porri
Fusarium Basal Rot	Fungus	Fusarium Oxysporum F. Sp. Cepae
Translucent Scale	Regular High Relative Humidity, and High Temperature (i.e. >90° F)	
Downy Mildew	Fungus	Peronospora Destructor

Fig. 2.   Common Onion Crop Diseases (a) Botrytis Blight (b) Botrytis Neck Rot (c) Purple Blotch (d) Fusarium Basal Rot (e) Translucent Scale (f) Downy Mildew.

The remaining article is arranged as follows: Sections II describes onion disease (purple blotch). Section III presents and summarizes related work. The state of art models for disease detection is discussed in Section IV. Section V explains what knowledge is needed to analyze the proposed model. Section VI gives a dataset description. The implementation details, performance metrics, and experimental results are described in Section VII. Lastly, Section VIII presents the outlines of the article's findings.

II.   PURPLE BLOTCH DISEASE

Younger leaves are more vulnerable than older leaves. Water-soaked lesions with a white center are the first signs of the disease. Lesions edges turn brown to purple the leaves above and below the lesions turn yellow. Dark brown to black concentric rings form over time inside the lesions. The fungus sporulates in these regions. Lesions can girdle the leaf as the disease progresses causing it to collapse and die as shown in Fig. 3. Seed stalks exhibit similar symptoms and infected stalks may collapse resulting in shriveled seed growth. The most common route for bulb infection is through the neck. If the fungus infects the bulb the infected area appears bright yellow at first, but then turns a distinctive red wine color.
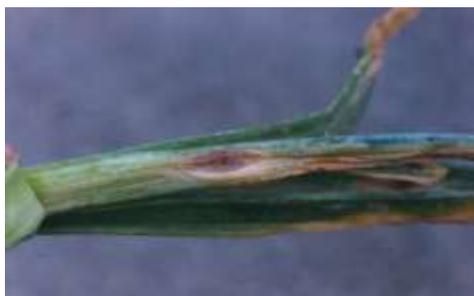

Fig. 3.   Onion Disease (Purple Blotch) [27].

a) Circumstances for Disease Expansion: In leaf debris and cull piles the fungus survives the winter as mycelium. During humid nights leaf wetness cycles exceed more than 12 hours spores are developed. When the morning dew evaporates spores become airborne and spread to vulnerable onion tissues. It takes 1-4 days for signs to appear after infection. During extended periods of leaf wetness, disease production is at its peak.

b) Overcome: A broad-spectrum protection fungicide spray program can provide good protection before infection. Reduce disease growth by reducing leaf wetness using surface irrigation rather than sprinkler irrigation, good field drainage, and proper plant spacing. Rotating onions to unrelated crops for many years can also help to minimize disease.

III.   RELATED WORK

Nihar et al. [28] suggested a neural network-based system for detecting plant disease that will aid in the development of the agricultural system which can properly determine whether a plant is infected or healthy and has a 97.7% accuracy. This technology enables the user to detect diseases faster allowing them to take acceptable precautionary steps and save crops.

Sangeetha et al. [29] compare different methods of image processing (K-mean clustering, artificial neural network (ANN), support vector machine (SVM), and fuzzy logic) then examine and summarize crop disease types. This study provides a thorough explanation of the machine learning (ML) models that were used to identify various agricultural diseases.

Liu et al. [30] presented a kiwi RGB-NIR-D dataset containing multi-modality associated images of kiwi fruits in farms. They suggested a unique method for using RGB-D sensors to combine associated near-infrared reflectance (NIR) and RGB images with deep learning (DL) techniques for fruit recognition. Using two modalities of aligned RGB and NIR images from the dataset faster region-based CNN is designed and deployed for kiwi fruit recognition.

Kim et al. [31] have proposed an image-built automated field monitoring method. Based on a weekly supervised learning approach the deep learning model was trained to identify crop disease. Using the field monitoring method the model was trained using captured onion crop field images, 6 groups including the sign of the disease were classified. As an ideal disease symptom localization threshold, 60% of the extreme value in the class activation map was calculated. The efficiency of recognizing disease symptoms was evaluated by intersection over union (IoU) using the mean average precision (mAP) metric.

Sharma et al. [32] enhanced decision-making with convolutional neural network (CNN) for various paddy crop diseases to prevent disease at initial stages and avoidance of mass loss in yield productivity. Paddy crop conditions are very fatal and can have serious effects on crops if initial treatment is not taken into account.

Karthik et al. [33] have used two distinct profound architectures to diagnose tomato leaf infection. To learn essential characteristics for classification the first architecture

relates to residual learning. On top of the residual deep network (RDN) the second architecture applies an attention function. Experiments were carried out over 3 diseases namely leaf mold, late blight, and early blight from the plant village dataset. Using the attention mechanism, the suggested work exploited the features acquired by CNN at different handling hierarchies and accomplished a whole accuracy of 98% in the 5-fold cross-validation on the validation sets.

Pattnaik et al. [34] suggested a transfer learning system for the classification of pests in tomato plants based on an existing deep CNN structure. The database for the analysis was obtained from 859 images in 10 groups from online sources. A comprehensive contrast of the performance of 15 existing CNN models has been offered. The test results showed that 88.83% of the maximum classification accuracy was achieved with the DenseNet169 model.

Francis et al. [35] offered a detailed summary of the prevalent applications of DL and computer vision (CV) techniques in the field of farming demonstrating the need to recognize and classify diseases using a dataset of leaf images. It is proposed to clarify its working principle with a novel classification system. A new collection containing gradient images is created using the multi-space image reconstruction input. Through the original and reconstructed images, high-level semantic structures are mined by convolutional layers that are convolutional and depth-wise separable. Finally, for classification SoftMax was used. The hyper-parameters and the expense of computation are calculated mathematically providing researchers with an insight innovation. The output of the system is compared and estimated with the related works on the openly accessible dataset of apple leaf images.

In this article, we have presented a methodology to detect onion disease (purple blotch) by using CNN classification. The feasibility of our approach has been demonstrated by using a dataset of 1000 images for healthy and infected onion crops. Two different types of experiments were taken using an improved InceptionV3 for classification with different batch sizes. Experiment trials used 70% of the dataset for training, 25% of the dataset for testing, and 5% of the dataset for validation. The proposed method recognizes the disease with a classification accuracy of 77.05% and 85.47% for different batch sizes.

IV. THE PROPOSED METHOD

The proposed method uses CNN to classify purple blotch disease in an onion crop. Finding an optimal architecture for CNN is one of the main challenges. Therefore we have used transfer learning where a model established for a mission is reprocessed as the beginning point for another task [36]. We have used a pre-trained CNN for recognition. This model is a functional model with multiple layers. The architecture of this model is shown in Fig. 4. This model is used to analyze the images and reliably detect them.

A. Input Layer

In image processing, it usually represents the image's pixel matrix. The parameter determines the dimension of the image (244×244×3) and due to data limitations, we augment our dataset by randomly generating images, adjusting the zoom and shear parameters of the original images. All images are duplicated five times. These copies are then fed into a pre-trained model, which is described further down.

B. Conv2D + ReLU Layer

Convolution is a linear procedure consisting of the multiplication between a two-dimensional (2D) weight array (filters) and an input data array of a weight set. We have 3 layers in the suggested design with zero padding and a filter of size 3×3. ReLU is abbreviated as a rectified linear unit that relates to the feature of non-saturating activation. It eliminates unwanted values from an activation map efficiently by setting them to nil. Without influencing the receptive fields of the convolution layer it rises the nonlinear assets of the resolution function and the whole network as specified in equation (1).

$$\text{ReLU} : f(x) = \max(0; x) \tag{1}$$

C. Pooling Layer

The pooling layer performs down-sampling feature maps by summarizing the presence of characteristics in the feature map patches. There are two kinds of pooling approaches max pooling and average pooling. We have used average pooling in the suggested design to determine the extreme value for each function map in each patch. The pooling average is set to 2×2 with 2 strides.
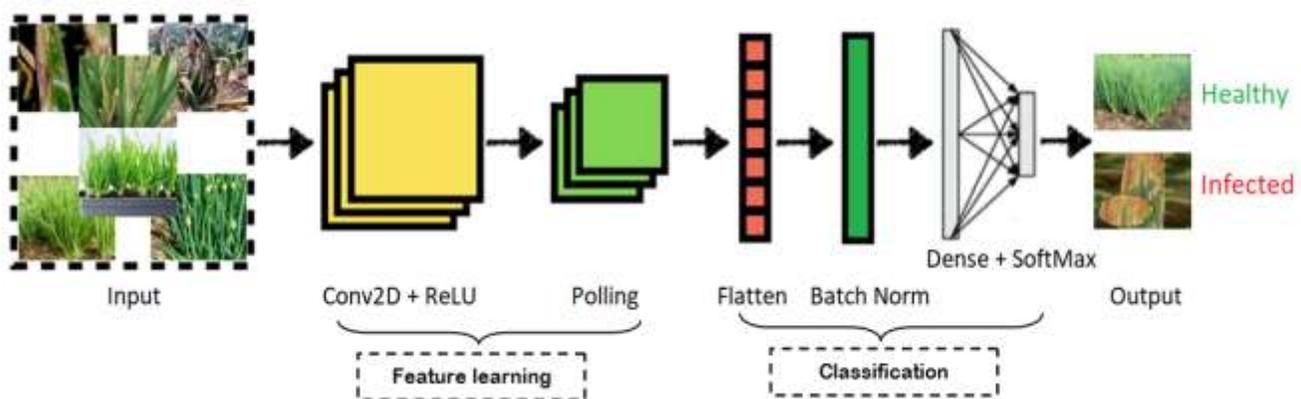

Fig. 4. CNN Architecture.

D. Flattening Layer

Flattening reduces file size by combining all apparent layers into the ground layer or transforms the data into a 1-dimensional array that is passed to the fully connected layer. To create a particular long feature vector we have dense the output of the convolutional layers.

E. Batch Normalization Layer

The batch normalization layer allows each layer to learn more individually. Learning becomes more rapid when batch normalization is utilized it can also be employed as a regularization to prevent model over-fitting. To standardize the outputs/ inputs, the layer is added to the model.

F. Dense + SoftMax Layer

The dense layer enhances an exciting non-linearity property, so any math function can be modeled. Though, they are always restricted in the logic that we still get the same output vector for the same input vector. On similar data, they can't identify recurrence in time or generate dissimilar responses. SoftMax layer is used to deliberate the feasible values of the preceding layer activation function. The values can be represented in two classes '0' and '1'. The logistic regression model is typically used for binary classification, although the SoftMax classification method is used for multi-classification. The classifier SoftMax is essentially an exponential normalized function as shown in equation (2).

$$\sigma(\hat{z}) = \frac{e^{zi}}{\sum_{j=1}^{K} e^{zj}} \tag{2}$$

Where

σ = Softmax

$\hat{z}$ = Input vector

e^{zi} = Standard input vector exponential function

K = Number of multi-class classifier classes

e^{zj} = Standard output vector exponential function

G. Output Layer

The output layer manages of generating the result. In a neural network, there must always be one output layer. The output layer receives the inputs from the layers above it executes the calculations using its neurons and then computes the output.

V. MODEL CONSIDERATION

One of the main goals of this research is to achieve state of art classification results using diverse transfer learning (DTL) models to both compensate for the small size of the trial data and to speed up the training procedure so that on modest hardware this can be achieved practically. These prototypes are all conveniently accessible as part of the Keras API and individually support TL in the form of support for pre-application to the model of the ImageNet weight.

A. Densenet201

DenseNet201 (dense convolutional network) is a 201 layers deep CNN that accepts an image input size of 224×224 [37]. DenseNet201 is a ResNet enhancement that requires dense connections between layers. In a feed-forward fashion, it ties each layer to every other layer. DensNet201 has $\frac{L(L+1)}{2}$ straight connections unlike typical CNN with L layers that have L connections. DenseNet has been able to boost efficiency compared to conventional networks by raising the necessity for computing, dropping the number of parameters, facilitating the reuse of features, and improving the propagation of features.

B. Inception-ResNet-V2

Inception-ResNet-V2 is a trained CNN on over a million ImageNet dataset images [38]. It is a cross procedure that incorporates the structure of residual connection and inception. The model takes a 299×299 dimension image and a list of approximate class probabilities in its output. The benefits of Inception-Resnet-V2 are the conversion of inception units to residual inception chunks and the adding of a new inception unit (Inception-A) to the stem module.

C. VGG16 and VGG19

VGG16 and VGG19 [39] are CNN architectures aimed to attain high accuracy in significant image identification applications with very narrow convolution filters (3×3). The depth [40] of the max-pooling, convolution, and fully connected layers vary between two implementations: 16 layers in VGG16 and 19 layers in VGG19.

D. MobileNet-V2

MobileNetV2 is an enhanced version of MobileNetV1 [41] it is consists of 54 layers and has an image input size of 224×224×3. Its key feature is that it uses two 1D convolutions with 2 kernels rather than having a single kernel 2D convolution. This implies that less storage, parameters, small and efficient model are needed for training. Two types of blocks can be distinguished: the first is a residual block with a stride of 1 and the other is blocked with a stride of 2 for downscaling. There are layers for each block: the first layer is ReLU6 1×1 convolution, the second layer is the deep convolution and another 1×1 convolution is the third layer but deprived of any non-linearity.

E. ResNet50

The ResNet was developed by executing a method of skip connections among layers known as residual learning to avoid the disappearing gradient problem characteristic in deep neural networks [42]. This design results in a network that trains extra effectively enable the design of deeper networks that have a positive effect on the accuracy of the model. ResNet50 is such a system implementing residual learning with 50 layers.

F. InceptionV3

The goal of InceptionV3 was to maximize the usage of computing assets within the network by raising the network's depth and width while retaining constant computing operations [43]. To define an enhanced network structure with avoided connections that are used as a building block the designers of this network coined the word 'Inception modules'. By assembling with max-pooling layers, [44] this initial

module is spatially repeated to minimize dimensionality to a reasonable smooth for calculation.

G. Xception

The "extreme" type of Inception model was developed by Google Inc [45]. Xception is consisting of 71 deep layers. Only depthwise detachable convolution layers are used in the Xception architecture.

VI. DATASET DESCRIPTION

On our image dataset, we applied our offerings for instinctive binary classification. The following section gives the steps of our contribution in detail.

A. Dataset

This present work introduces an image dataset that contains onion crop images. Under the guidance of Dr. Imtiaz A. Nizamani, Department of Plant Protection, Sindh Agriculture University, Tandojam, Pakistan images of onion crops were gathered from Tandojam, Sindh onion fields for this study. In this case, there are a total of 1000 images (png and jpeg format) with 600 images of healthy crops and 400 images of infected crops with 'purple blotch disease' which affects onion. However, when evaluating the onion crop will only reveal whether it is healthy or unhealthy. When it is unhealthy the onion crop will display four stages (low, medium, high, and fully infected). The crop becomes highly unstable, leaf length diminishes, flower excision occurs, and so on. To protect the land the sick crop should be separated from the healthy crop. Fig. 5 illustrates an example of the dataset used in this research where (a) shows infected onion images (b) shows healthy onion images. For splitting of data, we used in this experiment 70% of the dataset for training, 25% of the dataset for testing, and 5% of the dataset for validation. We guarantee that the images selected for validation are not used throughout training to accomplish the binary classification task. Furthermore, we perceived that our database is unprovoked certainly 40% of the images represent the infected onion class. To overwhelmed this issue by using data augmentation we re-sampled our dataset. We produced two new images with various augmentation techniques from each single input image. The total quantity of images was therefore increased by two times in the healthy and infected onion classes.

B. Data Augmentation

After dataset splitting, and pre-processing for the training process data augmentation is used it helps to prevent the possibility of over-fitting. The image of the onion was resized to 224×224. To prevent distortion the black background of 224×224 pixels was applied to the images with various pixel ratios to accomplish a comprehensive transformation. Low-contrast images of the whole onion crop were omitted. To make the requisite variety to help CNN's generalization competencies, data augmentation is mandatory. In particular, the images are arbitrarily rotated by a maximum of 10° and randomly moved vertically or horizontally in either direction by an all-out of 20 pixels. In this way, CNN acquires to be flexible in the variation of position and orientation alteration. Besides, geometric transformations such as rotations, re-scaling, shears, shifts, flips, zooms are the methods we have used, shown in Table VI.

TABLE VI.   DATA AUGMENTATION

Argument	Explanation
Rescaling	Rescaling images from integers 0-255 to float 0-1
Rotation	Degree variety of the arbitrary rotation
Shift	Shifts vertically or horizontally by (10%) and is the fraction of given dimension
Shear	Permit being sheared the angle in an anti-clockwise
Zoom	Permits the image to be "zoom-in" or "zoom-out"
Flip	Permitted the input to be flipped vertically or horizontally throughout the training procedure
Fill	The neighboring pixel value is repeated and chosen for all the vacant values

C. Data Limitation

Some drawbacks are faced by the dataset which has to be listed. Infected onion crops are incorporated into a relatively small sample. Moreover, the onion field with severe symptoms may be used to generate this sample analysis which is required. In the existing dataset crops with minor symptoms are absent which is due to the strategy of people with onion fields who have slight disease indications.

D. Classification and Training Dataset

When data pre-processing, data augmentation, and data splitting techniques are done our training onion dataset size is increased and arranged for the feature extraction phase with the model proposed to extract the necessary and applicable features. To construct vectorized feature maps the features extracted from the suggested model are compressed together. To classify the image into the corresponding groups the generated feature vector is transferred to a multi-layer perceptron. Lastly, on test images using the training model the efficiency of the proposed method is assessed. Each experiment is replicated three times and we report the average performance.

VII. EXPERIMENTAL SETUP

Our experiments have been carried out based on the following criteria for automatic binary classification established on our image dataset. All images in the dataset have been redimensioned to 224×224 pixels. We have set the batch size 8 and 16, the epochs are 500 to train our model. The samples are 155 and 90 respectively, for the training and validation. For optimization adam β1=0.89, β2=0.98, and learning rates are set at 10-5 and reduced to 10-6. Besides, we used decay weight to reduce our model over-fitting. The regularizers are supplied from Keras. The performance was specific to GPUs. Dell Inspiron Core(TM) i5-8250U CPU (8 CPU's) of 1.6 GHz, 16 GB Ram running on a Microsoft Windows 10 Professional (64-bit) is used in the implementation of the proposed model. The authors used Anaconda Navigator (Jupyter Notebook) version 3.7.4 to perform their research Keras, TensorFlow, OpenCV, NumPy, sklearrn are used for simulation as a backend for deep learning.
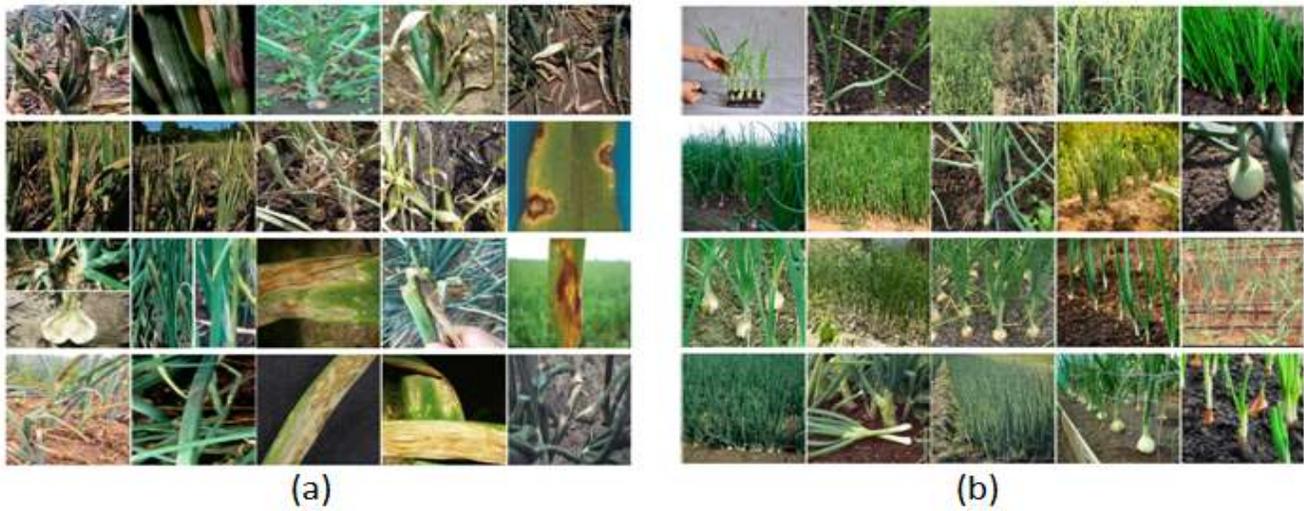
Fig. 5. Sample of Images used in Research (a) Infected Onion, (b) Healthy Onion.

A. Performance Metrics

The evaluation metrics that were used to measure the model's classification efficiency are explained in this section. For this, we use confusion matrix-based metrics [46]. Accuracy, recall, precision, F1-score, specificity, and sensitivity are examples of these metrics. We require the count of the following quantities to evaluate these measures: true positive, false positive, true negative, and false negative.

1) Accuracy: The ratio of correctly predicted predictions to the total number of predictions shown in equation (3).

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FN+FP} \tag{3}$$

2) Precision: The ratio of true positive prediction compared to overall positive predictions shown in equation (4).

$$\text{Precision} = \frac{TP}{TP+FP} \tag{4}$$

3) Recall: The ratio of a true positive prediction made by the suggested model to the total number of predictions made shown in equation (5).

$$\text{Recall} = \frac{TP}{TP+FN} \tag{5}$$

4) F1-Score: The harmonic mean of recall and precision is shown in equation (6).

$$\text{F1} - \text{Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{6}$$

5) Sensitivity: Measure the fraction of true positives correctly recognized shown in equation (7).

$$\text{Sensitivity} = \frac{TP}{TP+FN} \tag{7}$$

6) Specificity: (Also called TN rate) trials are performed on the fraction of true negatives properly-recognized shown in equation (8).

$$\text{Specificity} = \frac{TN}{TN+FP} \tag{8}$$

7) Confusion matrix: It is the evaluation of the model's results. It contrasts the true and predicted values in terms of true positive, false positive, true negative, and false negative shown in equation (9).

$$\begin{bmatrix} TP & FP \\ FN & TN \end{bmatrix} \tag{9}$$

- True Positive (TP): True positive predictions are those that were initially positive and were also predicted to be positive by the AI model.

- False Positive (FP): False positive predictions are those that were initially negative but were predicted as positive by the AI model.

- True Negative (TN): True negative predictions are those that were initially negative and were predicted as negative by the AI model.

- False Negative (FN): False-negative predictions are those that were initially positive but were predicted as negative by the AI model.

B. Experimental Results

Our research aimed to classify purple blotch disease from onion crop images using an improved InceptionV3 classifier we present the result for the binary classification in this section. Besides, several experiments are carried out on the onion crop disease (purple blotch) dataset to verify the efficiency and robustness of the suggested model. The findings are shown separately in Fig. 6 and Fig. 7 using the accuracy, loss, and receiver operating characteristic (ROC) curve of training and validation. The experiment is detailed described in Table VII. It indicates the overall outcome of the first trial batch size 8 and 500 Epochs. Accuracy and loss are calculated each after 100 epochs. We present the two most used parameters of computer vision and deep learning before discussing these results: Train curve is considered from the training dataset that gives an idea of how well the model is learning, while the validation curve is considered from a hold-out validation dataset that gives an idea of how well the model

is generalizing. Besides, the loss of validation and training is defined as a summing up of the errors made in the training sets of each case. The loss is not a fraction in comparison to accuracy. To summarize, the model that simplifies well, this is neither under-fitting nor over-fitting. The confusion matrix also displays a comprehensive illustration of what occurs to images later classification. We have observed that the accuracy increases until the value is 83.48% for the train and test accuracy from epoch 0 to epoch 7. The accuracy begins to be stable after epoch 10 were 87.01% and 84.42% for training and testing data respectively. In either the fast-growing interval from epoch 0 to epoch 32 where the loss is 2.98% or in the other interval where the decrease is slow and converges to 1.76% a good fit can be found for the loss curve of train results. As shown in the confusion matrix in Table IX, the model was able to classify 80 images as infected and 1 image as healthy for the infected onion class. The model with an improved InceptionV3 could predict 14 images as healthy and 27 images as infected for the healthy onion class. Similarly, Table VIII shows the overall result of the second trial with batch size 16 and 500 epochs. Likewise, accuracy and loss are

calculated each after 100 epochs. We have observed that the accuracy improves until the value is 88.98% for the training and testing from 0 epoch to 45 epoch. The accuracy is stable at epoch 38 where it is equivalent to 87.79% for training data and 86.87% for testing data. The loss curve of train data can be seen as a good match in the steadily rising interval between 0 and 100 epoch where the loss is 1.388% at the other time the decrease is gradual and converges to 1.14% at the end converges to 0.756%. As illustrated, confusion matrix in Table IX. The model can classify 78 images as infected and 1 image as healthy for the infected onion class. It can be shown that the InceptionV3 model was capable of predicting 22 images as healthy and 16 images as infected for the healthy onion class.

In the last, a comparison is made between batch sizes 8 and 16 as shown in Table IX predicted results are shown in Fig. 8. Generally, the outcomes of this trial show that the model pays complete courtesy to the features of the spot disease and achieves excellent onion disease recognition efficiency and it is very clear that batch size 16 provides approximate results.
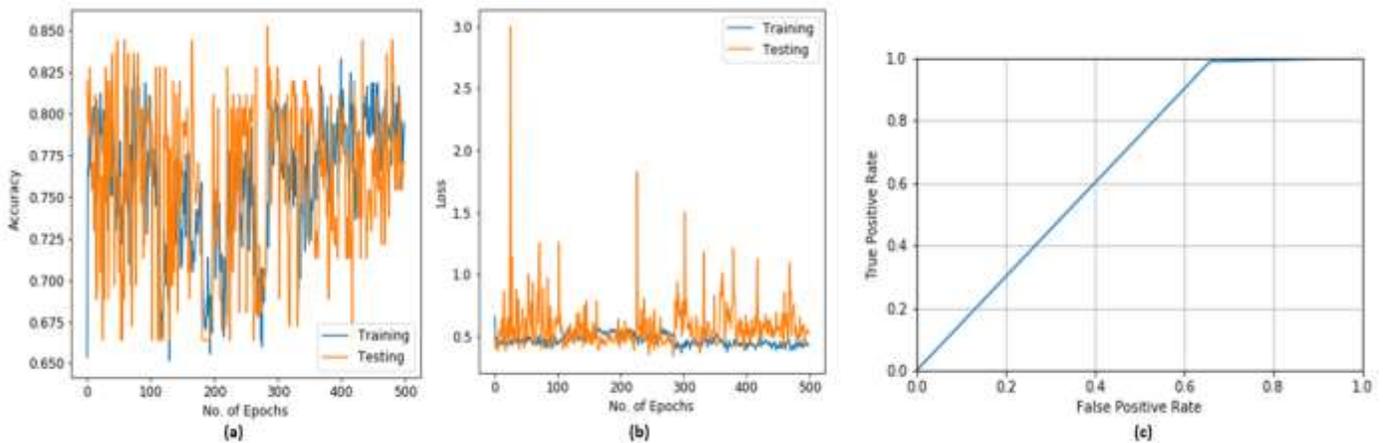

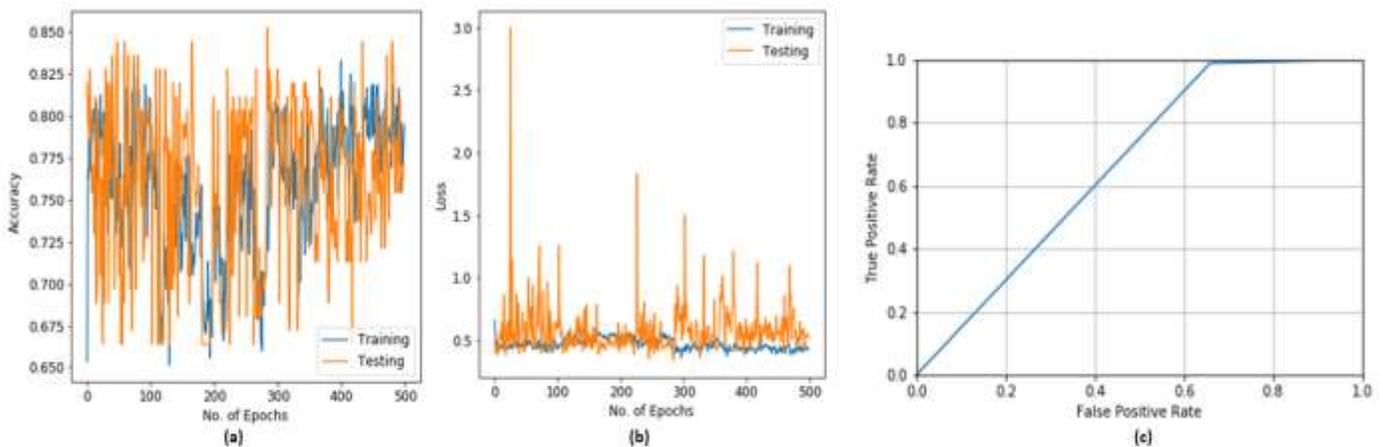Fig. 6.   CNN Model with Batch Size 8 (a) Accuracy, (b) Loss, (c) ROC Curve.

Fig. 7.   CNN Model with Batch Size 16 (a) Accuracy, (b) Loss, (c) ROC Curve.

TABLE VII.   RESULTS OF EXPERIMENT NO. 1

After	Accuracy/Loss
100 epochs	78.68% / 0.6275
200 epochs	81.15% / 0.4401
300 epochs	82.79% / 0.5441
400 epochs	78.68% / 0.5717
500 epochs	77.05% / 0.5197

TABLE VIII.   RESULTS OF EXPERIMENT NO. 2

After	Accuracy/Loss
100 epochs	85.47% / 0.4114
200 epochs	87.18% / 0.4993
300 epochs	84.62% / 0.3868
400 epochs	84.62% / 0.6148
500 epochs	85.47% / 0.4337

TABLE IX.   COMPARISON OF DIFFERENT BATCH SIZES

Batch Size	8	16
Confusion Matrix	[[14 27] [1 80]]	[[22 16] [1 78]]
Train Accuracy	79.38%	87.31%
Accuracy	77.05%	85.47%
Specificity	34.14%	57.89%
Sensitivity	98.76%	98.73%
Precision	74.76%	82.97%
Recall	98.76%	98.737%
F1-score	85.106	90.17

VIII. CONCLUSION

Alternaria porri is the fungus that causes purple blotch on onion crops. Leek, garlic, and chives are also affected by this fungus. On onion crops, spores germinate and create a thin water-soaked spot that turns brown. On several other crops, the oval-shaped lesion enlarges turns purplish, and forms the target spot appearance that Alternaria is known for (like early tomato blight). A yellow zone could encircle the margin. During wet weather, the surface of the lesion may be coated in brown to black masses of fungal spores. Infections in plants are a significant risk to worldwide food supplies. Extreme diseases in plants result in annual agricultural yield losses. The identification of diseases in plants at an initial stage is therefore very necessary for the avoidance of such dramatic losses in the future. This paper demonstrates the technical probability of deep learning to allow automatic infection analysis by image classification using the CNN approach. A new approach was discussed in this paper to use deep learning methods to spontaneously identify and detect crop disease from an image. The model established was able to differentiate between healthy and infected crops which can be diagnosed visually. The complete process was defined from the selection of images used for validation and training to the augmentation of images lastly, the deep CNN training procedure. We summarized the final results and concluded that through deep learning detection, segmentation, and classification our improved InceptionV3 achieves the highest precision as well as accuracy, F1-score, and recall. A deep CNN is accomplished to identify onion crops with a classification accuracy of 77.05% for batch size 8 and 85.47% for batch size 16 using our dataset of onion crops. Ongoing work aims to create a complete framework for crop monitoring. Besides, the performance can be improved by using a large dataset for biomedical image segmentation more advanced feature extraction techniques based on deep learning will be developed.
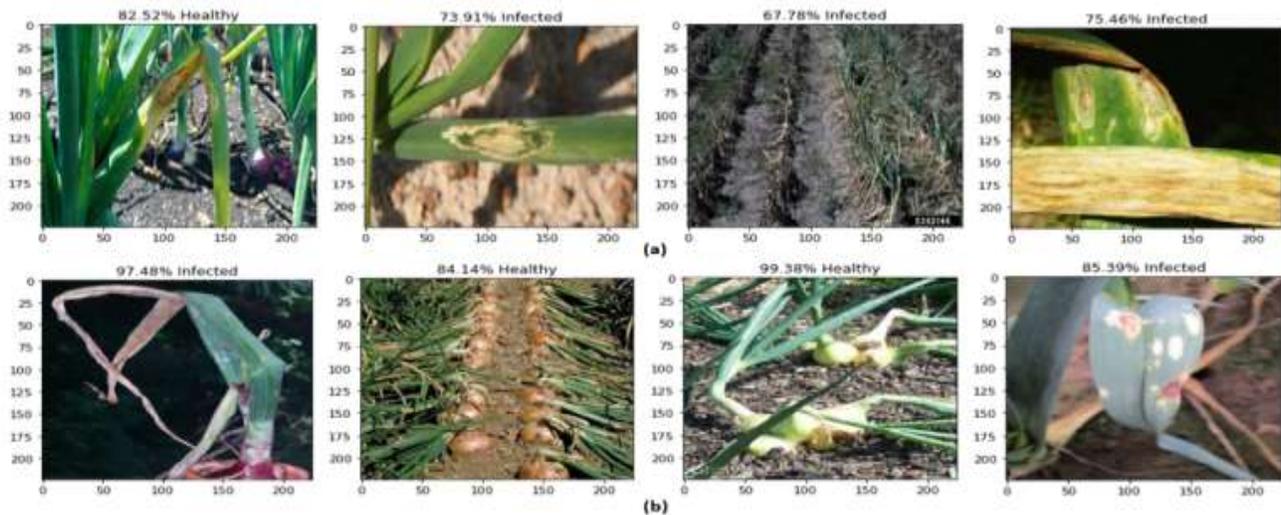

Fig. 8.   Predicted Results from (a) Batch Size 8, (b) Batch Size 16.

IX. FUTURE DIRECTIONS

1) Diseases and pests manifest themselves in different ways at different stages of growth. As a result pest and disease images should be differentiated with greater caution diseases, and pests of the same class should be divided according to the growth cycle as a rule. The division of the dataset will be improved in the future.

2) The next step will be to collect a huge number of high-quality images of diverse types of diseases and pests refine and adapt the model and extend it to other crops/plants to improve the practicability.

3) The disease detection efficiency of the enhanced InceptionV3 model will be improved even further. The goal of early diagnosis and detection of onion diseases can be accomplished by capturing lesions in real-time detection.

ACKNOWLEDGMENT

REFERENCES

[1] M. Jhuria, A. Kumar, and R. Borse, "Image processing for smart farming: Detection of disease and fruit grading," in *2013 IEEE Second International Conference on Image Information Processing (ICIIP-2013)*, 2013, pp. 521–526.

[2] S. Hena, L. Jingdong, A. Rehman, and O. Zhang, "A comparative analysis of agricultural development and modernization between China and Pakistan," *Int. J. Adv. Appl. Sci.*, vol. 6, no. 4, pp. 81–94, 2019.

[3] Y. Liu, A. Amin, S. F. Rasool, and Q. U. Zaman, "The Role of Agriculture and Foreign Remittances in Mitigating Rural Poverty: Empirical Evidence from Pakistan," *Risk Manag. Healthc. Policy*, vol. 13, p. 13, 2020.

[4] M. R. Faiz *et al.*, "Identification and prioritization of issues in growing and marketing vegetables in Punjab Province of Pakistan," *J. Innov. Sci.*, vol. 6, no. 1, pp. 54–59, 2020.

[5] A. Maqbool, S. A. Javeed, and A. Adeel, "Marketing of Carrot: A Case Study of Punjab Province, Pakistan."

[6] S. Abbas and Z. A. Mayo, "Impact of temperature and rainfall on rice production in Punjab, Pakistan," *Environ. Dev. Sustain.*, pp. 1–23, 2020.

[7] M. U. Shahbaz, T. Mukhtar, and N. Begum, "Biochemical and serological characterization of Ralstonia solanacearum associated with chilli seeds from Pakistan," *Int. J. Agric. Biol.*, vol. 17, no. 1, 2015.

[8] A. Hussain, M. A. Khathian, S. Ullah, and S. Ullah, "Increasing Crop Profitability through Adoption of Ridge Planting of Rice Crop in Punjab, and Wheat & Banana Crops in Sindh Province of Pakistan," *Proc. Pakistan Acad. Sci. B. Life Environ. Sci.*, vol. 56, no. 4, pp. 25–38, 2019.

[9] Hyesun Jang "Low Onion Prices of Major Asian Players, New Sources for European Markets" 2020. [Online]. Available: https://www.tridge.com/stories/low-onions-prices-of-major-asian-players-new-sources-for-european-markets

[10] C. B. T. Pal and G. C. Jadeja, "Microwave-assisted deep eutectic solvent extraction of phenolic antioxidants from onion (Allium cepa L.) peel: a Box–Behnken design approach for optimization," *J. Food Sci. Technol.*, vol. 56, no. 9, pp. 4211–4223, 2019.

[11] A. Nigussie, Y. Kuma, A. Adisu, T. Alemu, and K. Desalegn, "Onion production for income generation in small scale irrigation users agropastoral households of Ethiopia," *J. Hortic.*, pp. 1–5, 2015.

[12] L. Rajendran, "SCIENTOMETRIC ANALYSIS OF ONION (ALLIUM CEPA L) DURING 1980-2019: A STUDY BASED ON CAB DIRECT."

[13] G. Ekşi, A. M. G. Özkan, and M. Koyuncu, "Garlic and onions: An eastern tale," *J. Ethnopharmacol.*, p. 112675, 2020.

[14] N. B. Rathod and R. N. K. A. S. Karle, "Effect of organic, inorganic and integrated nutrient management on yield attributes, production of onion and soil properties, under vertisol of Maharashtra," *IJCS*, vol. 8, no. 1, pp. 937–940, 2020.

[15] C. Long, K. Hammer, and Z. Li, "The Central Asiatic region of cultivated plants," *Genet. Resour. Crop Evol.*, pp. 1–17, 2020.

[16] G. Messina, J. M. Peña, M. Vizzari, and G. Modica, "A Comparison of UAV and Satellites Multispectral Imagery in Monitoring Onion Crop. An Application in the 'Cipolla Rossa di Tropea'(Italy)," *Remote Sens.*, vol. 12, no. 20, p. 3424, 2020.

[17] H. Tamiru Geneti, "The response of onion (allium cepa l.) to applied water levels under pot planting at mehoni, raya valley of Ethiopia." Hawassa University, 2020.

[18] J. L. Brewster, *Onions and other vegetable alliums*, vol. 15. CABI, 2008.

[19] K. P. S. Kumar, D. Bhowmik, B. Chiranjib, and P. Tiwari, "Allium cepa: A traditional medicinal herb and its health benefits," *J. Chem. Pharm. Res.*, vol. 2, no. 1, pp. 283–291, 2010.

[20] S. Brankovic *et al.*, "Comparison of the hypotensive and bradycardic activity of ginkgo, garlic, and onion extracts," *Clin. Exp. Hypertens.*, vol. 33, no. 2, pp. 95–99, 2011.

[21] A. Kaur, T. G. Singh, S. Dhiman, S. Arora, and R. Babbar, "NOVEL HERBS USED IN COSMETICS FOR SKIN AND HAIR CARE: A REVIEW," *Plant Arch.*, vol. 20, no. 1, pp. 3784–3793, 2020.

[22] R. A. Baloch *et al.*, "Economic analysis of onion (Allium cepa L.) production and marketing in District Awaran, Balochistan," *Econ. Anal.*, vol. 5, no. 24, 2014.

[23] H. D. Lohano and F. M. Mari, "Spatial price linkages in regional onion markets of Pakistan," *J. Agric. Soc. Sci.*, vol. 1, pp. 318–321, 2005.

[24] M. A. Khokhar, K. M. Khokhar, and J. M. Khan, "WATER REQUIREMENT OF ONION CROP IN PAKISTAN."

[25] L. Black, K. Conn, B. Gabor, J. Kao, and J. Lutton, "Onion Disease Guide," *Seminis*, p. 71, 2012.

[26] S. Walker, N. Goldberg, and C. Cramer, "Onion Diseases in New Mexico," p. 12, 2014.

[27] L. Black, K. Conn, B. Gabor, J. Kao, and J. Lutton, "Onion Disease Guide," *Semin. grow Forw.*, p. 71, 2012.

[28] F. Nihar, N. N. Khanom, S. S. Hassan, and A. K. Das, "Plant Disease Detection through the Implementation of Diversified and Modified Neural Network Algorithms," *J. Eng. Adv.*, vol. 2, no. 1, pp. 48–57, 2021.

[29] S. K. B. Sangeetha, M. Sudha, R. Balamanigandan, and V. P. G. Pushparathi, "Comparison of Crop Disease Detection Methods - An intensive analysis," vol. 58, no. 2, pp. 10540–10546, 2021.

[30] Z. Liu *et al.*, "Improved Kiwifruit Detection Using Pre-Trained VGG16 with RGB and NIR Information Fusion," *IEEE Access*, vol. 8, no. January, pp. 2327–2336, 2020.

[31] W.-S. Kim, D.-H. Lee, and Y.-J. Kim, "Machine vision-based automatic disease symptom detection of onion downy mildew," *Comput. Electron. Agric.*, vol. 168, p. 105099, 2020.

[32] R. Sharma, S. Das, M. K. Gourisaria, S. S. Rautaray, and M. Pandey, "A Model for Prediction of Paddy Crop Disease Using CNN," in *Progress in Computing, Analytics and Networking*, Springer, 2020, pp. 533–543.

[33] R. Karthik, M. Hariharan, S. Anand, P. Mathikshara, A. Johnson, and R. Menaka, "Attention embedded residual CNN for disease detection in tomato leaves," *Appl. Soft Comput.*, vol. 86, p. 105933, 2020.

[34] G. Pattnaik, V. K. Shrivastava, and K. Parvathi, "Transfer Learning-Based Framework for Classification of Pest in Tomato Plants," *Appl. Artif. Intell.*, vol. 34, no. 13, pp. 981–993, 2020.

[35] M. Francis and C. Deisy, "Mathematical and Visual Understanding of a Deep Learning Model Towards m-Agriculture for Disease Diagnosis," *Arch. Comput. Methods Eng.*, pp. 1–17, 2020.

[36] W. Koehrsen, "Transfer Learning with Convolutional Neural Networks in PyTorch" 2018. [Online]. Available: https://towardsdatascience.com/transfer-learning-with-convolutional-neural-networks-in-%09pytorch-dd09190245ce.

[37] G. Huang, Z. Liu and N. Van Der Maaten, Densely Connected Convolutional Networks, The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.

[38] C. Szegedy, S. Ioffe, V. Vanhoucke and A. Alemi, Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning, arXiv:1602.07261, 2016.

[39] K. Simonyan and A. Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition arXiv preprint arXiv: 14091556, 2014.

[40] Q. Zhang, H. Wang, S. W. Yoon, D. Won, and K. Srihari, Lung Nodule Diagnosis on 3D Computed Tomography Images Using Deep Convolutional Neural Networks, Procedia Manufacturing, Vol. 39, pp. 363-370, 2019.

[41] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov and L-C. Chen, MobileNetV2: Inverted Residuals and Linear Bottlenecks, The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4510-4520, 2018.

[42] K. He, X. Zhang, S.Ren and J. Sunet, Deep Residual Learning for Image Recognition, Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778, 2016.

[43] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke and A. Rabinovich, Going Deeper with Convolutions, Technical report, 2014.

[44] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens and Z. Wojna, Rethinking the Inception Architecture for Computer Vision, arXiv preprint arXiv:1512.00567, 2015.

[45] F. Chollet, Xception: Deep Learning with Depthwise Separable Convolutions, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 21-26 July 2017.

[46] M. A. Zaki, S. Narejo, S. Zai, U. Zaki, Z. Altaf, and N. U Din, "Detection of nCoV-19 from Hybrid Dataset of CXR Images using Deep Convolutional Neural Network," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 12, pp. 699–707, 2020. *doi: 10.14569/ijacsa.2020.0111281.*